DO INFERIOR TEMPORAL CORTEX NEURONS ENCODE SHAPE BY ACTING AS FOURIER DESCRIPTOR FILTERS?

Thomas D. Albright and Charles G. Gross
The Salk Institute for Biological Studies, La Jolla, CA 92037 and Dept. Psychology, Princeton Univ., Princeton, NJ, 08544, USA

INTRODUCTION

Inferior temporal (IT) cortex is crucial for the perception of shape. Removal of this area produces severe impairments in learning and recognizing shapes while leaving basic sensory capacities, such as visual acuity, intact (2). Many IT neurons respond selectively to complex shapes and patterns (1,3,4). Often this selectivity is independent of the orientation, size or location of the shape.

How do IT cells extract information about overall shape? One possibility is that they do so with a method of describing boundary curvature known as the method of Fourier Descriptors (5), a procedure used in computer pattern recognition systems. This method depends first on determining the boundary orientation function for the shape, that is, the tangent angle of the shape's boundary measured at regular intervals around the perimeter. Then the boundary orientation function is expanded in a Fourier series. Each term in the Fourier expansion is associated with a particular frequency, amplitude and phase and is known as a Fourier Descriptor. Any closed boundary or shape is fully described by its set of Fourier Descriptors. The inverse transform of a single Fourier Descriptor uniquely determines a plane closed boundary having a specific number of lobes (frequency), lobe indentation (amplitude), and phase (orientation). (We call these shapes determined by a single Fourier Descriptor, "FD stimuli".) This method of describing shapes is independent of both the position and size of the stimulus. Thus, Fourier Descriptors are an efficient method for representing shapes.

In an earlier paper (4) we proposed that IT neurons might code or represent shape by acting as detectors or filters for Fourier descriptors. This hypothesis yields several testable predictions. First, at least some IT cells should be selective for specific FD stimuli. Second, this selectivity should cover an adequate range of parameters for coding an arbitrary shape. Third, this selectivity should remain invariant over variations in stimulus size, exact retinal location and contrast. A fourth, and particularly stringent test is that the response of an IT cell to a complex shape should be predictable from the cell's responses to FD stimuli.

In a previous paper (4) we reported tests of the first three of these predictions and found that about half of the visually responsive IT cell sampled were systematically tuned to the frequency of the FD stimuli. For the majority of the cells tested, this selectivity for frequency remained constant over changes in the size, retinal location and contrast of the FD stimuli. Furthermore, many IT cells were also selective for the amplitude and phase of the FD stimuli. Thus, many IT cells met the first three requirements of our hypothesis that IT cells represent shape by acting as filters for Fourier Descriptors.

In the present paper we report the fourth and critical test, namely that the response of an IT neuron to a complex closed boundary can be predicted from its response to simple FD stimuli. Our results were clearly incompatible with the hypothesis that single IT units act as linear filters for boundary curvature. Furthermore, we failed to find evidence for any coherent non-linear coding scheme involving Fourier Descriptors.

METHOD

Two *Macaca fascicularis* were each recorded from repeatedly over a 4-8 week period. Details of surgical preparation and recording procedures have been previously reported (1). All surgical procedures were performed under strictly aseptic conditions using barbiturate anesthesia. During recording sessions animals were immobilized and ventilated with N_2O/O_2. At least 2 days separated successive recording sessions. Varnish coated tungsten microelectrodes were used to record extracellular potentials

from single isolated neurons. All recording sites were located within the central portion of cytoarchitectonic area TE.

After each unit was isolated its receptive field was estimated using simple stimuli such as slits, edges and bars that were viewed monocularly through the eye contralateral to the recording site. A set of "simple FD" stimuli was produced, as previously described (3,4), by inverting boundary angle functions defined as a set of pure sinusoids. A second set of stimuli, "compound FDs", was produced by inverting boundary angle functions defined as paired linear combinations of sinusoids. Each compound FD was thus composed of two frequencies and the set included all possible combinations of the frequencies that constituted the set of simple Fds. Finally, a third set of stimuli, "sectioned Fds", was produced by sectioning all members of the simple FD set into "halves" and "quarters". In order to interpret the effects of these sectioned stimuli, we computed the Fourier transform of the boundary angle function for each. All FD stimuli were filled (white on black) boundaries of uniform intensity, equated for total area, and positioned at the center of gaze.

If a neuron was found to be selective for frequency of simple FD stimuli, a subset of compound Fds was chosen to include stimuli composed of frequencies that elicited combinations of strong and weak responses when presented as simple FDs. A subset of sectioned FDs was also chosen to include, typically, halved versions of the frequencies eliciting maximal and minimal responses. The data were analyzed by comparing responses predicted in accordance with a "linear FD filter" hypothesis with the patterns of selectivity actually obtained from single IT neurons. According to this hypothesis, addition of a ineffective frequency (little or no response) to a highly effective one (maximal response) should have a negligible effect on the response to the resultant compound FD stimulus. The predicted response to any compound FD stimulus was approximated from the response to the most effective simple FD in the pair. The responses to sectioned FDs were also predicted according to the simple linear FD filter hypothesis: the predicted response was approximated from the response to the most effective of the individual frequency components comprising the sectioned FD.

RESULTS

FD frequency selectivity

One hundred thirty-five visually responsive IT neurons were sampled along eleven penetrations in two animals. Forty-eight percent (65/135) of the sample failed to respond to any of the simple FDs. Ninety-four percent (66/70) of the responsive cells responded differentially as a function of frequency, while the remainder failed to show any FD frequency selectivity.

Responses to compound FDs

Twenty-four cells were tested with a set of compound FDs following initial assessment of FD frequency selectivity. Examples of data from two different cells are shown in Figs 1 and 2.

The cell in Fig 1 responded maximally to f3 and responses declined monotonically with increasing frequency (Fig 1A). This cell was tested with compound FDs produced by combinations of f3 with two ineffective frequencies, f11 and f32. A compound FD derived from f11 and f32 was also used. The responses to these compound FDs are shown in Fig 1B along with the predictions from our linear FD filter hypothesis, which anticipates that the response to a combination of frequency components will be equal to the response to the most effective component frequency in the pair. The responses to the compound FDs differ significantly from our predictions: although ineffective alone, addition of f11 or f32 actually reduced sensitivity to the f3 frequency component.

The cell in Fig 2 responded strongly to f32, while responses to the other seven frequencies were negligible (Fig 2A). This cell was tested with two compound FDs produced by combining f32 with two ineffective frequencies, f3 and f8. The responses to these compound FDs are plotted in Fig 2B along with the linear FD filter predictions. As can be seen, the responses to each of these compound FDs are considerably smaller than our predictions; the cell seems incapable of detecting the presence of the f32 component in the company of non-optimal frequencies.
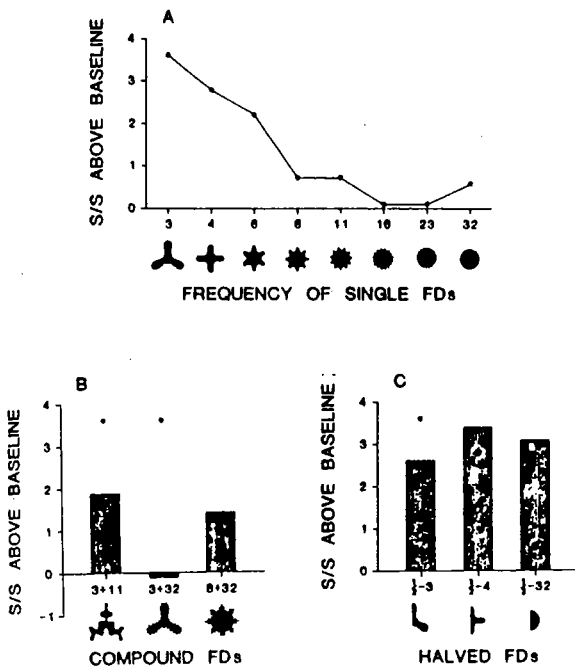
Fig 1: A: Selectivity of a single IT neuron for FD frequency. B: Actual (bars) and predicted responses (dots) of the same neuron for compound FDs. C: Actual and predicted responses to sectioned (halved) FDs.
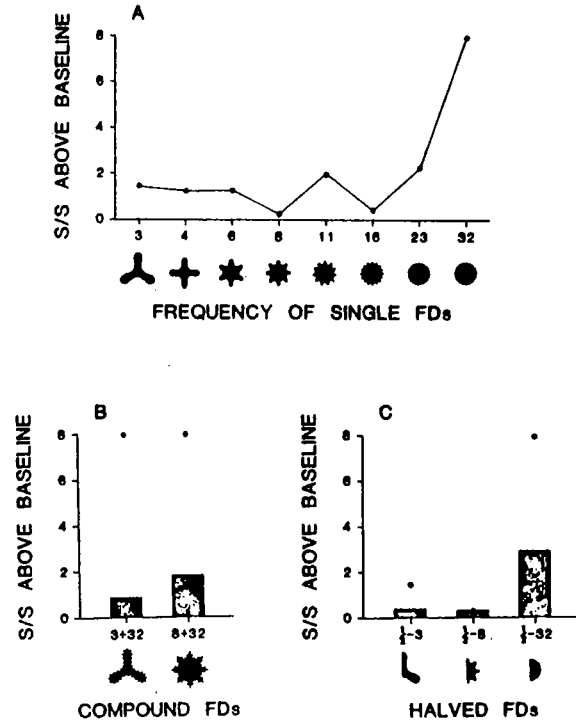
Fig 2: A: Selectivity of a single IT neuron for FD frequency. B: Actual and predicted responses of the same neuron for compound FDs. C: Actual and predicted responses to sectioned (halved) FDs.

The overall relationship between our linear FD filter predictions and the responses of single IT neurons was quantified by a simple correlation between predicted (based on responses to simple FDs) and actual responses to compound FDs. The distribution of these coefficients is centered at zero, indicating a complete failure to support our linear FD filter hypothesis.

Responses to sectioned FDs

Sixteen cells were tested with a set of sectioned FDs following initial assessment of FD frequency selectivity. Examples of data from two such cells are shown in Figs 1C and 2C. The frequency selectivity data for the same two neurons may be seen in Figs 1A and 2A and their compound FD responses may be seen in Figs 1B and 2B.

The first example, which responded maximally to f3, was tested with sectioned FDs derived from f3 as well as from f4, which also elicited a good response, and f32, an ineffective FD frequency. The responses to these sectioned FDs are plotted in Fig 1C along with the predictions based upon the frequency content of the sectioned FD. According to this hypothesis, the cell in Fig 1 should respond best to sectioned FD frequency 3 (of the three tested), since this is the stimulus that still contains the largest complement of frequency 3. The responses to the sectioned FDs differ significantly from our predictions. In fact, responses to the three sectioned FDs were nearly identical, suggesting that neither frequency spectra (in accord with our conclusions from the compound FD manipulations reported above) nor shape "elements" (local aspects of the boundary contour) are relevant parameters for this particular IT neuron, despite notable selectivity for low frequency simple FDs.

The second example, which responded to f32 with negligible responses to the other simple FDs, was tested with sectioned

FDs derived from f32 as well as from f3 and f8. The responses to these sectioned FDs are plotted in Fig 2C along with the predictions from our two alternative hypotheses. According to the linear FD filter hypotheses, the cell should respond best to sectioned FD frequency 32 (of the three tested), since this is the stimulus that now contains the largest complement of frequency 32. Although behavior of this cell under these conditions was marginally consistent with the linear FD filter hypothesis (responses being strongest to sectioned FD f32), the same cell failed the compound FD test (Fig 2B). An alternative interpretation is suggested by the fact that responses to the three sectioned FDs used on this cell mirrored the pattern of selectivity for simple FDs. This cell's behavior is thus consistent with some type of selectivity for local aspects of boundary curvature (however vaguely defined).

The overall relationship between our linear FD filter predictions and the responses of single IT neurons was quantified by calculation of a simple correlation between predicted (based on responses to simple FDs) and actual responses to sectioned FDs. The distribution of these coefficients is centered at zero, indicating a complete failure to support our linear FD filter hypothesis.

Non-linear FD filters?

Our data permit us to reject the possibility that IT neurons act as linear FD filters. The possibility of some non-linear mechanism is much more difficult to dismiss. We have observed, however, no consistent behavior across our sample of IT neurons. Although inconclusive, this fact greatly lessens the plausibility of shape encoding by IT neurons acting as non-linear FD filters.

DISCUSSION

We have performed the most direct and stringent test of the hypothesis that IT neurons encode complex shapes by acting as filters for Fourier Descriptors. To satisfy this test we required that the response of an IT neuron to a complex shape be predictable from knowledge of 1) the cell's selectivity for simple FD frequency and 2) the FD frequency composition of the complex shape. Using this criterion we failed to find any support for a linear filtering process. We also failed to find evidence for any consistent non-linear encoding scheme, although this possibility cannot be conclusively ruled-out, however unlikely.

The remarkable fact remains, nonetheless, that many IT neurons are selective for the frequency of simple FD stimuli. These affectively neutral stimuli thus provide a powerful means of characterizing and quantifying the shape selectivity of IT neurons (3,4). They have also proved useful for demonstrating important IT traits such as position and size invariance and for studying response habituation. The neural interactions that confer selectivity for FD stimuli and the underlying reasons for this property remain mysteries ripe for further investigation.

REFERENCES

1.   Desimone, R., Albright, T.D., Gross, C.G. and Bruce, C. (1985) *J. Neurosci.* 4:2051-2062.

2.   Gross, C.G. (1972) In: R. Jung (ed), *Handbook of Sensory Physiology*, Vol. VIII/3B. Springer-Verlag, Berlin, pp. 451-482.

3.   Gross, C.G., Desimone, R., Albright, T.D. and Schwartz, E.L. In: C. Chagas, R. Gattass, and C.G. Gross (eds), *Pattern Recognition Mechanisms*, Springer-Verlag, New York, pp. 179-201.

4.   Schwartz, E.L., Desimone, R., Albright, T.D. and Gross, C.G. (1982) *Proc. Nat. Acad. Sci. USA.* **80**:5776-5778.

5.   Zahn, C.T and Roskies, R.Z. (1972) *IEEE Trans. Computers.* C-21:269-281.